

Metodika pro automatizovanou tvorbu slovníku znakového jazyka

Jakub Kanis, Marek Hruží, Pavel Campř

Katedra kybernetiky, Fakulta aplikovaných věd,
Západočeská univerzita v Plzni, jkanis, mhruz, campř@kky.zcu.cz

Anotace: Tento článek je věnován popisu metodiky, jejímž cílem je zrychlit a zjednodušit tvorbu slovníků znakového jazyka. Tato metodika je založena na využití pokročilých metod zpracování digitalizovaného obrazu. Ty umožňují například automatickou detekci hranic jednotlivých znaků ve video nahrávce obsahující posloupnost znaků a dále také automatickou kategorizaci znaků na základě jejich automaticky rozpoznávaných atributů.

Úvod

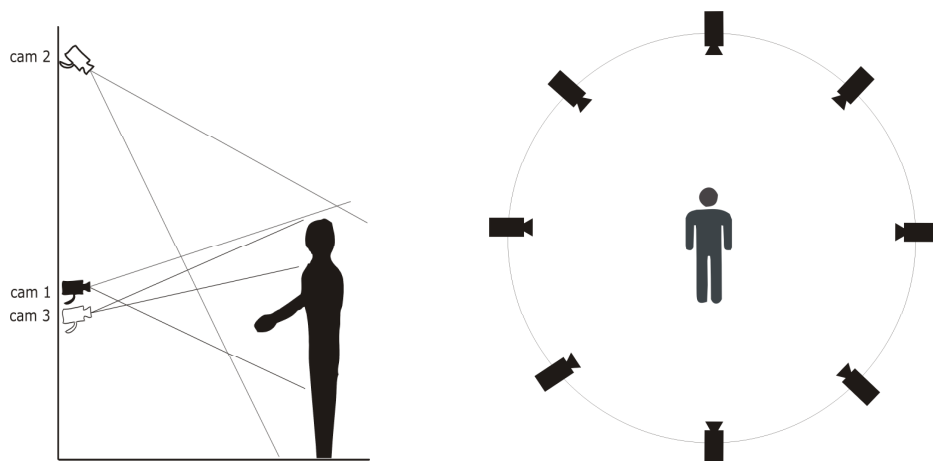
Slovník je důležitým zdrojem o daném jazyce. To platí především v případě znakových jazyků, kde neexistuje žádná oficiální psaná forma a slovníky a lexikony jsou jedním z nejdůležitějších zdrojů pro výuku a lingvistický výzkum znakových jazyků. Cílem našeho projektu je vytvořit on-line slovník znakového jazyka, který bude schopen splnit tyto funkce a bude tedy jak výkladový tak také překladový. Do slovníku budou zahrnuty stávající i nové (nově nahrané) znaky. V tomto článku představíme metodiku vhodnou pro automatizované zpracování nahraných znaků a jejich přidání do vytvářeného slovníku.

Metodika

Dále bude popsána metodika pro nahrávání nových znaků a jejich zpracování pro přidání do libovolného slovníku znakového jazyka.

Scéna

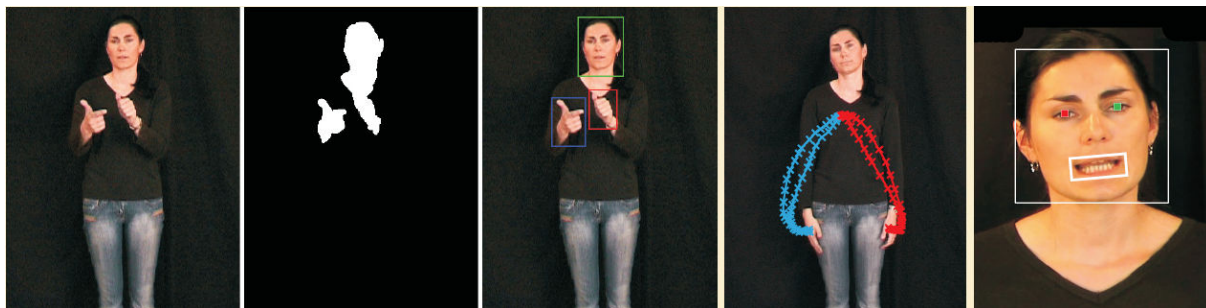
Scéna by měla mít jednobarevné pozadí, nejlépe černé barvy. Do obrazu žádné kamery by neměly zasahovat objekty jiné barvy, nejlépe by se v obraze neměly vyskytovat vůbec. **V žádném případě se nesmí v kameře objevit barva kůže jinde než na části lidského těla.** Scéna by tudíž měla být prázdná a nejlépe černá. Velikost nahrávacího prostoru by měla být úměrná velikosti lidského těla, přičemž je nutné brát v úvahu rozmanitost výšky a šířky u lidí. V záznamu musí být vidět celý základní znakovací prostor, který je zhruba vymezen pasem znakujícího, rozpaženými pažemi a temenem hlavy. Existují znaky, při kterých ruce opustí tento prostor. Sice jsou považovány za výjimky, ovšem scéna musí být připravena tak, aby bylo možné zachytit i tyto znaky. Bylo by vhodné scénu změřit a vytvořit její matematický model.



Obrázek 1: Ukázky možných konfigurací kamer

Kamery a světla

Je vhodné použít více kamer. Jednak jsou pak ukázky přehlednější pro člověka, kdy nahrávky nabudou trojrozměrný charakter. S více kamerami je rovněž ulehčena automatická detekce objektů. Největším problémem při automatickém zpracování obrazu je překryv nebo splynutí více objektů v jeden. V našem případě jde o ruce. Pak je vhodné zvolit různé pohledy tak, aby aspoň jeden poskytl objekty v nepřekrývajících se konfiguraci. **Platí pravidlo: čím více kamer tím lépe.** Měl by se tudíž volit takový počet kamer, aby byly dodrženy body z odstavce **Scéna** a zároveň jich bylo co nejvíc. Za samozřejmý se považuje pohled zepředu. Pohledy zpoza znakujícího nemají větší význam, protože znakující nemění během záznamu svou orientaci. Nejméně přínosný je pak pohled ze strany. V tomto pohledu dochází k častým překryvům, které se musí řešit pomocí jiného pohledu. Vhodný pohled je pohled ze shora, mírně před znakujícím. Tento pohled vidí většinou ruce, které jsou z předního pohledu v překryvu, nespojené. U tohoto pohledu se však vyskytuje problém umístění kamery. Zjednodušeně řečeno je pro tento pohled potřeba vysokého stropu. Pokud by se volily další pohledy, tak je vhodné nechávat kamery vysoko a měnit jejich polohu po kružnici s vhodným poloměrem a se středem v bodě kde stojí znakující (Obrázek 1). Důležitá informace o znaku je zachycena i ve tváři znakujícího. Proto stojí za zmínku možnost využít jednoho detailního pohledu na tvář znakujícího.



Obrázek 2: Ukázka záznamu a jeho zpracování

Co se týče parametrů jednotlivých kamer, zase lze uplatnit pravidlo čím více tím lépe. Kvalita dnešních HD kamer je daleko za potřebami automatického zpracování obrazu. Ovšem vždy je jednodušší kvalitu zmenšovat, nežli zvyšovat. Pro automatické zpracování je PAL (720px*576px) rozlišení postačující. Je to z důvodu velkého datového toku, který vzniká při zpracování videa. Hardware v PC není v dnešní době schopen zpracovat tak velké množství dat v rozumném čase. **Kamery by měli mít barevný záznam s konstantním počtem snímků za vteřinu.** Tyto vlastnosti jsou pro automatické zpracování velmi důležité. Druhá vlastnost je nesmírně důležitá pro synchronizaci více kamer. Doba uzávěrky a clona by měli být nastaveny tak, aby si jednotlivé pohledy odpovídali. Je důležité zachovat kvalitu barev, neměli by být moc „vyšedlé“, ale ani přesvícené. **Všechny automatické funkce kamer, musí být vypnuty.** Jako příklad uvedeme zaostření (focus). Záznam by neměl v průběhu měnit svou charakteristiku. Zoom kamery by měl být vypnutý. **Určitě nepoužívat digitální zoom,** který mění snímání obraz. Zoom může být nastaven tak, aby se nesnímalo zbytečně velký prostor, ale musí být zachován během celého nahrávání.

Počet a umístění světel by mělo být takové, aby nevznikaly nežádoucí stíny a neměnil se charakter přirozených barev (bílé světlo). Rovněž by scéna neměla být zaplavena světlem, aby jasy nebyli v kameře „přepálené“. Tento bod je nutné odzkoušet přímo „za běhu“. Většinou jsou dvě světla postačující.

Herec

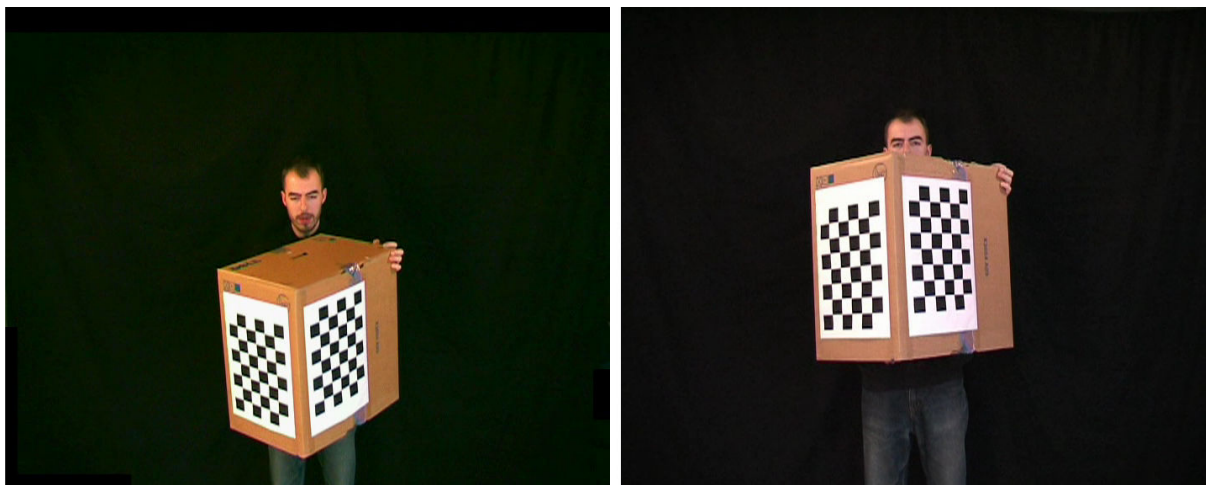
Herec by měl být oblečen v jednobarevném oblečení bez textury. Barva nemusí být černá, ale měla by být lehce odlišitelná od barvy kůže. Z RGB složek má nejmenší podíl v barvě kůže zelená a tak by barva oblečení měla mít zelenou jako dominantní barvu. V případě stejné barvy oblečení a pozadí, se ztrácí informace o zbytku těla, která může mít dobrý vliv na úspěšnost automatického rozpoznávání. Za zmínku stojí to, že momentálně implementované metody nevyužívají tuto informaci, tudíž tato podmínka není silná.

Herec by měl stát na nehybné značce, tak aby byl v pohledu zepředu ve středu obrazu. **Měla by být zvolena základní poloha herce,** tj. poloha kdy není žádný znak prováděn. Tato poloha musí být jednotná, nejlépe pro všechny herce a všechny záznamy. Námi používána poloha jsou ruce spuštěné podél těla. **Je dobré, aby v základní poloze bylo vidět obě ruce,** neboť to značně usnadní počáteční hledání objektů ve scéně.

Kalibrační vzor a klapka

Klapka je důležitá kvůli synchronizaci více kamer. Je nesmírně důležité, aby byla vidět ve všech kamerách a nejlépe byla doprovázena silným zvukovým impulsem. Alternativně se může zvolit jiný způsob synchronizace. Např. zatmění a opětovné rozsvícení v místnosti nebo vyslat ze scény na všechny kamery silný světelný signál (blesk, ale ne bodový). Tyto alternativy jsou vhodné pro automatický střih videa. Silné světelné změny jsou snadno detekovatelné, ovšem v jiných kamerách se můžou projevit jinak. **V každém případě musí být použita klapka.** Pokud by selhala automatická synchronizace, vždy musí být v záznamu klapka, aby synchronizaci mohl určit člověk.

Kalibrační vzor se používá za předpokladu, že bude zájem o 3D informaci ze scény. Je vhodné jej použít i jinak, neboť to nezabere moc času a pak přináší nesmírné zjednodušení při určování hloubky bodů ve scéně. Zároveň je jednodušší určit vzájemný vztah mezi kamerami a tak využívat jejich „spolupráci“. Jako kalibrační vzor se většinou používá šachovnice s černo-bílými poli (Obrázek 3). Tato šachovnice nesmí být koplanární (tj. nesmí být rovina), nejlépe by měla být v půli zlomena na dvě na sebe kolmé části. **Kalibrační vzor musí být celý vidět vždy minimálně v páru kamer** (tím jsme schopni určit vzájemný vztah mezi dvěma kamerami). V čem více kamerách je kalibrační vzor vidět, tím více vztahů můžeme vyvodit a tím přesnější je 3D rekonstrukce bodů scény. **Kalibrační vzor musí být změřen** a následně sestaven jeho matematický model.



Obrázek 3: Kalibrační vzor

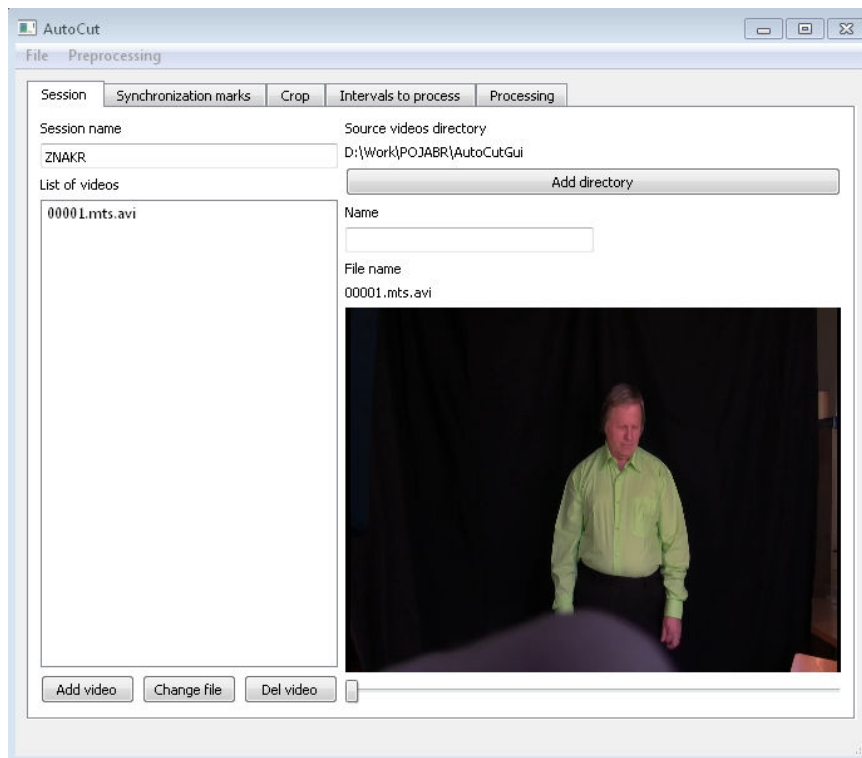
Postup natáčení

1. Zvolit vhodnou konfiguraci kamer a světel dle odstavce **Scéna a Kamery a světla**
2. Ujistit se, že kamery mají nastavené správné hodnoty parametrů, a že jsou vypnuta všechna automatická nastavení.
3. Postupně nastavit intenzitu světel a hodnoty parametrů kamer tak, aby obraz vyhovoval odstavci **Kamery a světla**.
4. Zapnout všechny kamery na záznam.
5. Provést synchronizaci kamer – klapka. Odstavec **Kalibrační vzor a klapka**.
6. Provést kalibraci kamer – kalibrační vzor. Odstavec **Kalibrační vzor a klapka**.
7. Postavit herce do základní polohy.
8. Může začít samotný záznam. Pokud dojde kdykoliv k přerušení záznamu, musí se opět splnit body 1. – 7.
9. Na závěr záznamu je vhodné udělat novou klapku.

AutoCut

Abychom urychlili zpracování pořízených video nahrávek, vytvořili jsme specializovaný program. Program byl vytvořen s ohledem na to, aby uživatele provedl celým zpracováním videa od předzpracování až k zveřejnění nahraných znaku ve slovníku. Prvním krokem při zpracování nahrávek je jejich převod do formátu

vhodného pro další zpracování. K tomu slouží uživateli jednoduchý dialog pro výběr zpracovávaných souborů a spuštění samotného předzpracování. Hlavní okno programu (viz Obrázek 4) má pět karet, které jsou v pořadí odpovídajícím obvyklému postupu při zpracování zvolených nahrávek. Na první kartě může uživatel přidat, změnit nebo smazat zdrojová videa (vytvořena při předzpracování). Na druhé kartě může být nastavená synchronizační značka, která je důležitá pro zpracování nahrávek pořízených z různých úhlů pohledu. Další karta slouží k nastavení ořezu a čtvrtá pro určení zpracovávaných intervalu. Poslední karta je určena k automatické detekci hranic znaku, nastavení modelu barvy pokožky a finálnímu generování jednotlivých znaku vhodných pro zveřejnění.



Obrázek 4: AutoCut

Závěr

V tomto článku byla představena ucelená metodika pro automatizovanou tvorbu slovníku znakového jazyka. Metodika se zabývá popisem postupu od nahrávání znaků až po jejich zpracování a přípravu k publikaci ve slovníku. Díky využití pokročilých metod automatického zpracování digitalizovaného obrazu, lze při dodržení této metodiky získat významnou časovou úsporu oproti současnému způsobu zpracování nahraných znaků (např. automatický versus ruční střih znaků).

Poděkování

Tato práce vznikla v rámci projektu CZ.1.07/2.2.00/07.0189 „Potlačení jazykové bariéry sluchově postižených studentů prostřednictvím automatického zpracování jazyka“ financovaného Evropským sociálním fondem a státním rozpočtem České republiky.